

## A new parameterization for estimating co-occurrence of interacting species

J. HARDIN WADDLE,<sup>1,6</sup> ROBERT M. DORAZIO,<sup>2,3</sup> SUSAN C. WALLS,<sup>3</sup> KENNETH G. RICE,<sup>3</sup> JEFF BEAUCHAMP,<sup>4</sup>  
MELINDA J. SCHUMAN,<sup>5</sup> AND FRANK J. MAZZOTTI<sup>4</sup>

<sup>1</sup>U.S. Geological Survey, National Wetlands Research Center, Lafayette, Louisiana 70506 USA

<sup>2</sup>University of Florida, Department of Statistics, Gainesville, Florida 32611 USA

<sup>3</sup>U.S. Geological Survey, Southeast Ecological Science Center, Gainesville, Florida 32653 USA

<sup>4</sup>University of Florida, Fort Lauderdale Research and Educational Center, Davie, Florida 33314 USA

<sup>5</sup>Conservancy of Southwest Florida, Naples, Florida 34102 USA

**Abstract.** Models currently used to estimate patterns of species co-occurrence while accounting for errors in detection of species can be difficult to fit when the effects of covariates on species occurrence probabilities are included. The source of the estimation problems is the particular parameterization used to specify species co-occurrence probability. We develop a new parameterization for estimating patterns of co-occurrence of interacting species that allows the effects of covariates to be specified quite naturally without estimation problems. In our model, the occurrence of one species is assumed to depend on the occurrence of another, but the occurrence of the second species is not assumed to depend on the presence of the first species. This pattern of co-occurrence, wherein one species is dominant and the other is subordinate, can be produced by several types of ecological interactions (predator–prey, parasitism, and so on).

A simulation study demonstrated that estimates of species occurrence probabilities were unbiased in samples of 50–100 locations and three surveys per location, provided species are easily detected (probability of detection  $\geq 0.5$ ). Higher sample sizes ( $>200$  locations) are needed to achieve unbiasedness when species are more difficult to detect. An analysis of data from treefrog surveys in southern Florida indicated that the occurrence of Cuban treefrogs, an invasive predator species, was highest near the point of its introduction and declined with distance from that location. Sites occupied by Cuban treefrogs were 9.0 times less likely to contain green treefrogs and 15.7 times less likely to contain squirrel treefrogs compared to sites without Cuban treefrogs. The detection probabilities of native treefrog species did not depend on the presence of Cuban treefrogs, suggesting that the native treefrog species are naive to the introduced species.

**Key words:** anurans; detectability; Florida, USA; hierarchical model; invasive species; predation; treefrogs.

### INTRODUCTION

Distinguishing patterns of co-occurrence between species is an important first step in identifying the potential for ecological interactions, as well as for understanding the role that such interactions may play in creating spatial heterogeneity in community composition. Competition, predation, and parasitism are all examples of ecological interactions in which one species is negatively affected by another. In communities where prey and predator species also potentially compete with one another, intraguild predation can influence the composition of species in complex ways (Polis et al. 1989). The outcome of these interactions can vary depending on the intensity of interspecific competition

(Chase et al. 2002). However, such negative interactions, either alone or in combination with one another, can yield patterns of co-occurrence wherein one species is present less often in local habitats occupied by the other species.

MacKenzie et al. (2004) developed a useful model for estimating probabilities of co-occurrence between pairs of species while accounting for errors in detection of each species. In addition, this model can be used to determine whether the presence of one species affects the detectability of the other species. In typical applications of this model, species occurrence and detection probabilities are specified as functions of observed covariates, such as habitat characteristics or sampling effort, that are thought to be informative of the spatial heterogeneity in occurrence and detection of each species. The effects of these covariates are formulated as logit-scale regression parameters. Unfortunately, MacKenzie et al. (2004) noted that numerical problems can occur when attempting to estimate these parameters due to restric-

Manuscript received 15 May 2009; revised 17 September 2009; accepted 21 September 2009. Corresponding Editor: J. Van Buskirk.

<sup>6</sup> E-mail: waddleh@usgs.gov

tions that must be imposed on their values. These restrictions stem from the particular parameterization of species co-occurrence and detection adopted by MacKenzie et al. (2004).

Here, we develop a new parameterization for estimating patterns of co-occurrence of interacting species. This parameterization allows the effects of covariates to be specified without the estimation problems described by MacKenzie et al. (2004). In our approach, the occurrence of one species is assumed to depend on the occurrence of another, but the occurrence of the second species is not assumed to depend on the presence of the first species. This pattern of co-occurrence, wherein one species is dominant and the other is subordinate, can be produced by several types of ecological interactions, as described earlier. In our model, the conditional probabilities of detection are treated similar to occurrence probabilities, that is, detection probabilities of the subordinate species are allowed to differ depending on whether or not a site is occupied by the dominant species.

Using simulated data, we show that our model of co-occurrence can provide unbiased estimates of its parameters. We illustrate our approach by analyzing patterns of co-occurrence of treefrog species in southern Florida, USA. The Cuban treefrog (*Osteopilus septentrionalis*), a nonnative, introduced species in Florida, is reported to be a voracious predator of native treefrog species (Meshaka et al. 2000, Smith 2005). Cuban treefrogs may also compete with native treefrogs for food owing to the high degree of dietary overlap among these species (Meshaka 2001). Although the Cuban treefrog is believed to have expanded its range in Florida primarily through human facilitation (Meshaka 1996), this species also appears to be invading less-traveled natural areas. Thus, Cuban treefrogs represent a potential threat to native treefrog species in these areas, but the ecological consequences of the spread of this invasive species are not known. We use our model of species co-occurrence to assess whether the occurrence probabilities of native treefrog species differ in the presence and absence of Cuban treefrogs while accounting for differences in occurrence associated with differences in habitat. We also explore the possibility that detection probabilities of native treefrog species differ between sites with and without Cuban treefrogs. Such differences could be induced by cryptic behavior of native species attempting to avoid predation by Cuban treefrogs.

#### A MODEL FOR ESTIMATING CO-OCCURRENCE OF INTERACTING SPECIES

We describe a statistical model for estimating patterns of co-occurrence between a single dominant species and one or more subordinate species when observations of all species are subject to imperfect detection. The model extends ideas developed for the analysis of single-species occurrence (MacKenzie et al. 2002) to the analysis of

data from multiple species. We formulate the co-occurrence of species in a way that honors our assumed interactions between species. Specifically, we consider situations where the occurrence of each subordinate species is assumed to depend on the occurrence of a dominant species, but the presence of the dominant species is independent of that of each subordinate. An example exists in predator–prey interactions wherein the mean occurrence of the prey species is reduced by the presence of the predator species but the mean occurrence of the predator is independent of the presence of any single prey species. For ease of presentation we describe the model using the terms predator and prey, instead of dominant and subordinate species, even though our model assumptions can be applied to many types of ecological interactions.

We develop the model using a hierarchical (or state-space) formulation that includes distinct components for species occurrence and species detection given occurrence. This approach allows us to build models of predator and prey species occurrences and their potential ecological determinants (e.g., habitat or other site-level covariates) independent of the effects of imperfect detection and sampling (Royle and Dorazio 2008). The relevance of these latter effects is specified in a second component of the model, the observation process, wherein detections of species are modeled conditional on latent parameters for predator and prey species occurrences. This hierarchical approach provides a generic framework for the construction of occupancy models, and these models can be fit to data using either classical (frequentist) or Bayesian methods (see Appendix).

#### Modeling species occurrences

We begin by describing a model of species occurrences for a single sample location. For the moment, we suppress the location subscript on model parameters. Let  $z^A$  and  $z^B$  denote a pair of binary random variables for the occupancy state of prey species  $A$  and predator species  $B$ , respectively (wherein  $z = 1$  indicates species presence and  $z = 0$  indicates species absence). Thus,  $(z^A, z^B) = (0, 1)$  corresponds to an observation wherein species  $A$  is absent and species  $B$  is present. For ease of presentation, we describe the model with only one prey species; however, any number of prey species can be included simply by adding parameters to the model for those species. This flexibility is attributed to the manner in which we model associations between predator and prey species occurrences.

We use three parameters to model the joint distribution of occupancy states of species  $A$  and  $B$  [i.e., to model  $(z^A, z^B)$ ]:

- 1)  $\psi^B = \Pr(z^B = 1) = \text{probability of occurrence (presence) of predator species } B$ ;
- 2)  $\psi^{A|B} = \Pr(z^A = 1 \mid z^B = 1) = \text{conditional probability of occurrence of prey species } A \text{ given that predator species } B \text{ is present;}$

3)  $\psi^{A|B} = \Pr(z^A = 1 | z^B = 0)$  = conditional probability of occurrence of prey species  $A$  given that predator species  $B$  is absent.

Using these parameters, the joint probability of the occupancy states of species  $A$  and  $B$  can be expressed as a function of their marginal and conditional probabilities of occurrence since  $\Pr(z^A, z^B) = \Pr(z^A | z^B) \Pr(z^B)$ . Therefore, we model the joint distribution of occupancy states of species  $A$  and  $B$  as a product of the following two Bernoulli processes:

$$z^B | \psi^B \sim \text{Bernoulli}(\psi^B) \leftarrow$$

$$z^A | z^B, \psi^{A|B}, \psi^{A|\bar{B}} \sim \text{Bernoulli}\left(z^B \psi^{A|B} + (1 - z^B) \psi^{A|\bar{B}}\right).$$

In this way, the occupancy state of species  $A$  depends explicitly on that of species  $B$ . Biologically speaking, this construction implies that prey species  $A$  is present with one probability ( $\psi^{A|B}$ ) in the presence of predator  $B$  and with another probability ( $\psi^{A|\bar{B}}$ ) in the absence of the predator. On the other hand, if the occurrence of species  $A$  does not depend on whether species  $B$  is present or absent, this alternative modeling assumption can be specified simply by assuming  $\psi^{A|B} = \psi^{A|\bar{B}}$ , that is, by assuming a restricted version of the model of predator–prey association.

#### Modeling species detections

As with other occupancy-based approaches, we assume that  $J > 1$  replicate observations are made independently at each of  $n$  sample locations and that the duration of sampling is sufficiently short that the occupancy state of a species is fixed at each location during the survey. Within-site replication provides the information needed to compute estimates of both species occurrence and species detection probabilities. Limits on duration of the survey satisfy the standard assumption of site-specific, demographic closure in occupancy state.

Let  $\mathbf{y} = (y_1, y_2, \dots, y_J)$  denote an encounter history composed of  $J$  binary observations that indicate whether a species was detected ( $y_j = 1$ ) or not detected ( $y_j = 0$ ) during the  $j$ th observation of the sample location. For example,  $\mathbf{y} = (0, 1, 0, 1, 0)$  represents the history of a species that was detected twice (during the second and fourth observations) in  $J = 5$  replicate surveys of the sample location. We denote the encounter histories of species  $A$  and  $B$  by  $\mathbf{y}^A$  and  $\mathbf{y}^B$ , respectively.

Our ability to detect a species at any site depends on whether that species is present at the site. Therefore, we specify a model of each species' encounter history conditional on the binary occupancy state (present or absent) of that species. Furthermore, we assume that the  $J$  observations are made independently, which is consistent with our sampling protocol. Such assumptions are common in models of species occurrence that account for imperfect detectability of animals (MacKenzie et al. 2006, Royle and Dorazio 2008).

Given these assumptions, each element of the encounter history of predator species  $B$  is modeled as follows:

$$y^B | z^B, p^B \sim \text{Bernoulli}(z^B p^B)$$

where  $p^B = \Pr(y^B = 1 | z^B = 1)$  denotes the conditional probability that species  $B$  is detected during a single observation of the sample location, given that this species is present. Thus, if the location is unoccupied by species  $B$  (i.e.,  $z^B = 0$ ), then  $\mathbf{y}^B = \mathbf{0}$  with probability 1; otherwise, species  $B$  is detected with probability  $p^B$  during each observation.

The detection probability of the prey species is assumed to depend on whether the predator species is present or absent. Such differences in detection could be attributed to a behavioral response of the prey, being more difficult to detect in the presence of the predator. Whatever the source, we include two different parameters for prey detection probabilities:

1)  $p^{A|B} = \Pr(y^A = 1 | z^A = 1, z^B = 1)$  = conditional probability that prey species  $A$  is detected during a single observation of the sample location, given that species  $A$  and predator species  $B$  are both present;

2)  $p^{A|\bar{B}} = \Pr(y^A = 1 | z^A = 1, z^B = 0)$  = conditional probability that prey species  $A$  is detected during a single observation of the sample location, given that species  $A$  is present and predator species  $B$  is absent.

Using these parameters, each element of the encounter history of species  $A$  is modeled as follows:

$$y^A | z^A, z^B, p^{A|B}, p^{A|\bar{B}} \sim \text{Bernoulli}\left(z^A \left\{ z^B p^{A|B} + (1 - z^B) p^{A|\bar{B}} \right\}\right).$$

Thus, if the location is unoccupied by species  $A$  (i.e.,  $z^A = 0$ ), then  $\mathbf{y}^A = \mathbf{0}$  with probability 1; otherwise, species  $A$  is detected with probability  $p^{A|B}$  or  $p^{A|\bar{B}}$  during each observation, depending on the presence or absence of the predator species  $B$ . Note that if the detection of prey species  $A$  does not depend on whether species  $B$  is present or absent, this alternative modeling assumption can be specified simply by assuming  $p^{A|B} = p^{A|\bar{B}}$ , that is, by assuming a restricted version of the model of predator–prey association.

#### Modeling effects of covariates on probabilities of species occurrence and detection

The occurrence of a prey species is unlikely to depend solely on whether its predator is present or absent at a particular sample location. Other factors, such as habitat characteristics, also may be important in determining the occurrence of the prey and the predator species. Moreover, these factors can differ between species. For example, in our survey of treefrogs (see *Example: treefrog species in southern Florida*) the mean occurrence of the predator species (Cuban treefrog) is expected to decline with distance from the point of introduction of this nonnative species. The occurrences

of both native and nonnative species may depend on a sample location's habitat: pine savanna (relatively dry) or swamp forest (relatively wet).

The detectability of a prey species also may depend on factors other than presence or absence of the predator. In our treefrog survey, we expect air temperature at the time and location of sampling to be associated with the probabilities of detection of both prey and predator species. Our expectation is that the native treefrog species were less active and harder to detect at lower temperatures than at higher temperatures.

Here we introduce notation to include the effects of site-level covariates on probabilities of species occurrence and the effects of site- or replicate-level covariates on conditional probabilities of species detection. We use the covariates described above as examples to illustrate our modeling approach. Let  $i$  index the sample location ( $i = 1, \dots, n$ ) and  $j$  index a replicate observation from each location ( $j = 1, \dots, J$ ). Given this notation,  $z_i^B$  denotes a binary random variable for occurrence of predator species  $B$  at location  $i$  and  $\psi_i^B$  denotes the probability of occurrence of species  $B$  at that location. The effects of longitude and habitat on  $\psi_i^B$  are formulated as logit-scale parameters as follows:

$$\text{logit}(\psi_i^B) = \beta_{01}^B \text{pine}_i + \beta_{02}^B (1 - \text{pine}_i) + \beta_1^B \text{lon}_i$$

where  $\text{pine}_i$  indicates the habitat (1 = pine savanna, 0 = swamp forest) and  $\text{lon}_i$  indicates the longitude of the  $i$ th sample location. The observed values of longitude are centered and scaled to have zero mean and unit variance; therefore, the parameters  $\beta_{01}^B$  and  $\beta_{02}^B$  can be interpreted as the logit-scale probabilities of occurrence of species  $B$  in pine savanna and swamp forest habitats, respectively, at the average longitude of the sample.

We specify a model of site-specific occurrence probability of prey species  $A$  similarly, except that we also must include its dependence on occupancy state of the predator. Let  $\psi_i^A$  denote the probability of occurrence of prey species  $A$  at sample location  $i$ . The effects of predator occupancy state and habitat on  $\psi_i^A$  are formulated as logit-scale parameters as follows:

$$\begin{aligned} \text{logit}(\psi_i^A) = & \beta_{01}^A z_i^B \text{pine}_i + \beta_{02}^A (1 - z_i^B) \text{pine}_i \\ & + \beta_{11}^A z_i^B (1 - \text{pine}_i) + \beta_{12}^A (1 - z_i^B) (1 - \text{pine}_i). \end{aligned}$$

Thus,  $\beta_{01}^A$  and  $\beta_{02}^A$  correspond to logit-scale parameterizations of the conditional probabilities of occurrence of species  $A$  with and without predators ( $\psi^{A|B}$  and  $\psi^{A|\bar{B}}$ , respectively) in pine savanna habitat. Similarly,  $\beta_{11}^A$  and  $\beta_{12}^A$  correspond to logit-scale parameterizations of these probabilities in swamp forest habitat.

Let  $p_{ij}^B$  denote the conditional probability of detection of predator species  $B$  during the  $j$ th replicate observation of sample location  $i$ . We model the effects of habitat and air temperature on  $p_{ij}^B$  on the logit-scale, as we did earlier for site-specific occurrence probability:

$$\text{logit}(p_{ij}^B) = \alpha_{01}^B \text{pine}_i + \alpha_{02}^B (1 - \text{pine}_i) + \alpha_1^B \text{temp}_{ij}$$

where  $\text{temp}_{ij}$  indicates the centered and scaled temperature observed during the  $j$ th observation of sample location  $i$ . Therefore, the parameters  $\alpha_{01}^B$  and  $\alpha_{02}^B$  can be interpreted as the logit-scale probabilities of detection of species  $B$  in pine savanna and swamp forest habitats, respectively, at the average temperature observed during the entire survey.

We formulate a model of the detection probabilities of prey species  $A$  similarly, except that we also include its dependence on the occupancy state of the predator at the  $i$ th sample location:

$$\text{logit}(p_{ij}^A) = \alpha_{01}^A z_i^B + \alpha_{02}^A (1 - z_i^B) + \alpha_1^A \text{temp}_{ij}.$$

Thus,  $\alpha_{01}^A$  and  $\alpha_{02}^A$  correspond to logit-scale parameterizations of the conditional probabilities of detection of species  $A$  with and without predators ( $p^{A|B}$  and  $p^{A|\bar{B}}$ , respectively) at the average temperature observed during the entire survey.

Our hierarchical model contains many latent parameters; however, the model can be fitted using either classical methods (i.e., maximizing the marginal likelihood obtained by removing latent parameters by integration [see Appendix]) or Bayesian methods (see Supplement).

## SIMULATION STUDY

In this section we describe a simulation study that was conducted to evaluate the performance of our model-based estimator of species co-occurrence. In the study, we considered the range of sample sizes (number of sample locations  $n$  and number of surveys per location  $J$ ), detection probabilities, and probabilities of occurrence and co-occurrence used in the simulations of MacKenzie et al. (2004). Doing so allowed us to compare our estimator of co-occurrence with that of MacKenzie et al. (2004); it also allowed us to make some recommendations regarding survey design.

### Methods

For each combination of sample sizes and model parameters, 1000 data sets were randomly generated using the assumptions of our model of co-occurrence and detection. As in the simulations of MacKenzie et al. (2004), we considered samples from  $n = 50, 100$ , or 200 locations and  $J = 3$  or 5 replicate surveys per location. For simplicity, the conditional probabilities of detection ( $p^B$ ,  $p^{A|B}$  and  $p^{A|\bar{B}}$ ) were assigned equal values of 0.214 or 0.5.

In the model of MacKenzie et al. (2004),  $\psi^{AB}$  denotes the joint probability of occurrence of two species ( $A$  and  $B$ ). In our model this probability is denoted using the product of conditional and marginal probabilities of species occurrence:  $\psi^{AB} = \psi^{A|B} \psi^B$ . The simulations of MacKenzie et al. (2004) included the following values of  $\psi^{AB}$ : 0.08, 0.24, 0.4, and 0.6125. The first two of these values were paired with *unconditional* probabilities of occurrence of  $\psi^A = \psi^B = 0.4$ ; the latter two values of  $\psi^{AB}$  were paired with higher occurrence probabilities,  $\psi^A = \psi^B = 0.6$ .



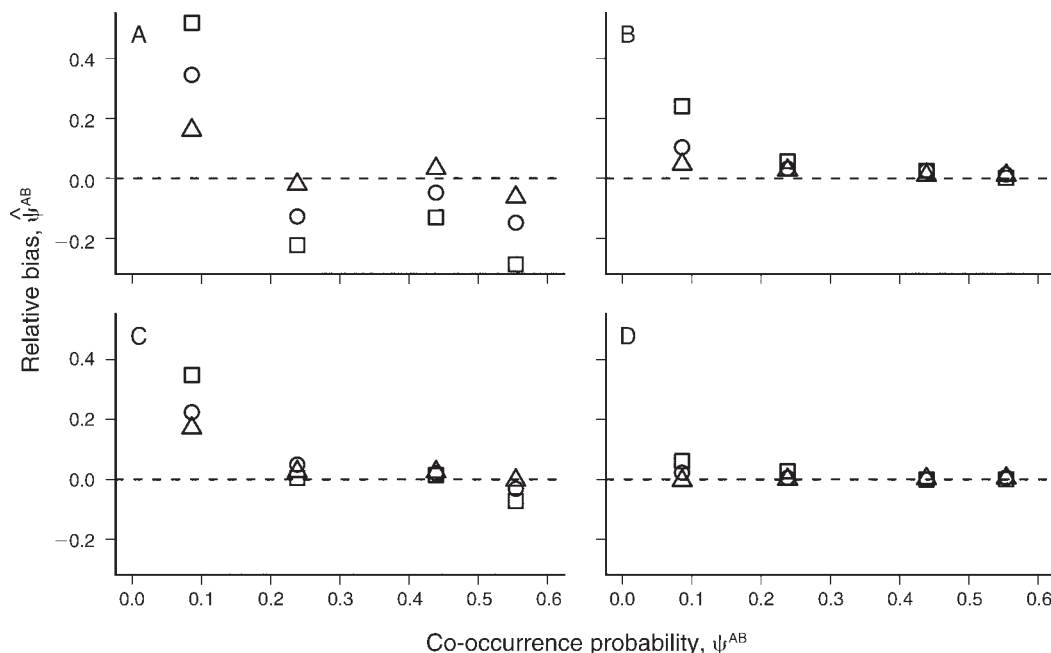


FIG. 1. Simulation-based approximations of relative (indicated by hat) bias of our estimator of the probability of co-occurrence  $\psi^{AB}$  between two species (A and B) as a function of the true value of  $\psi^{AB}$ . Symbols indicate number of sample locations  $n$  (square = 50, circle = 100, triangle = 200). Panels correspond to different combinations of detection probability and number of replicate surveys (A,  $p = 0.214$ ,  $J = 3$ ; B,  $p = 0.5$ ,  $J = 3$ ; C,  $p = 0.214$ ,  $J = 5$ ; D,  $p = 0.5$ ,  $J = 5$ ).

$\psi^B = 0.7$ . These combinations of occurrence and co-occurrence probabilities were assumed to reflect either a strong association or a strong disassociation between occurrences of species A and B.

In our simulations, we used the same unconditional probabilities of occurrence used by MacKenzie et al. (2004); however, our probabilities of co-occurrence  $\psi^{AB}$  differed slightly. To be specific, we selected values of  $\psi^{AB}$  needed to make the odds ratio of the conditional probabilities  $\psi^{A|B}$  and  $\psi^{A|\bar{B}}$  equal to either 4 or 1/4. Therefore, values of  $\psi^{AB}$  were selected so that the odds of occurrence of species A were four times higher (or four times lower) in the presence of species B than in the absence of this species. To summarize, we used the following four combinations of  $\{\psi^A, \psi^B, \psi^{AB}\}$  in our simulations:  $\{0.4, 0.4, 0.08610\}$ ,  $\{0.4, 0.4, 0.23837\}$ ,  $\{0.7, 0.7, 0.43885\}$ ,  $\{0.7, 0.7, 0.55397\}$ .

We computed maximum likelihood estimates of the parameters of our model ( $\psi^{A|B}$ ,  $\psi^{A|\bar{B}}$ ,  $\psi^B$ ,  $p^{A|B}$ ,  $p^{A|\bar{B}}$ ,  $p^B$ ) for each set of simulated data. These estimates were computed without assuming any restrictions on the parameters, just as in the simulations of MacKenzie et al. (2004). The relative bias of the maximum likelihood estimator of co-occurrence probability was approximated by expressing the difference between the average of the 1000 simulated estimates of  $\psi^{AB}$  (say,  $E(\hat{\psi}^{AB})$ ) and the data-generating value of  $\psi^{AB}$  as a proportion (i.e.,  $(E(\hat{\psi}^{AB}) - \psi^{AB})/\psi^{AB}$ ). We excluded simulations in which the observed information matrix was nearly singular (reciprocal condition number  $< 0.0001$ ) and therefore not invertible, as done by MacKenzie et al. (2004). These

cases typically were associated with estimates of occurrence parameters near their boundary values (0 or 1).

### Results

The results of our simulation study are similar to those reported by MacKenzie et al. (2004). Our model-based estimator of co-occurrence probability  $\psi^{AB}$  was unbiased in most of the simulations except those with low sample size ( $n = 50$  or  $100$  and  $J = 3$ ) and low detection probability ( $p = 0.214$ ; Fig. 1). In these cases the bias was highest at the lowest co-occurrence probability ( $\psi^{AB} = 0.08610$ ), which corresponds to conditional occurrence probabilities of  $\psi^{A|B} = 0.215$  (for species A in the presence of species B) and  $\psi^{A|\bar{B}} = 0.523$  (for species A in the absence of species B). At all co-occurrence probabilities, the bias decreased with increases in the number of sample locations, which has important implications for survey design (see Discussion).

### EXAMPLE: TREEFROG SPECIES IN SOUTHERN FLORIDA

#### Methods

During 2007–2008, we surveyed natural areas of southern Florida for the purpose of estimating patterns of co-occurrence between Cuban treefrogs and native treefrog species. Sampling was conducted at 107 sites across Picayune State Forest, Fakahatchee Strand Preserve State Park, Florida Panther National Wildlife Refuge, and Big Cypress National Preserve in Collier

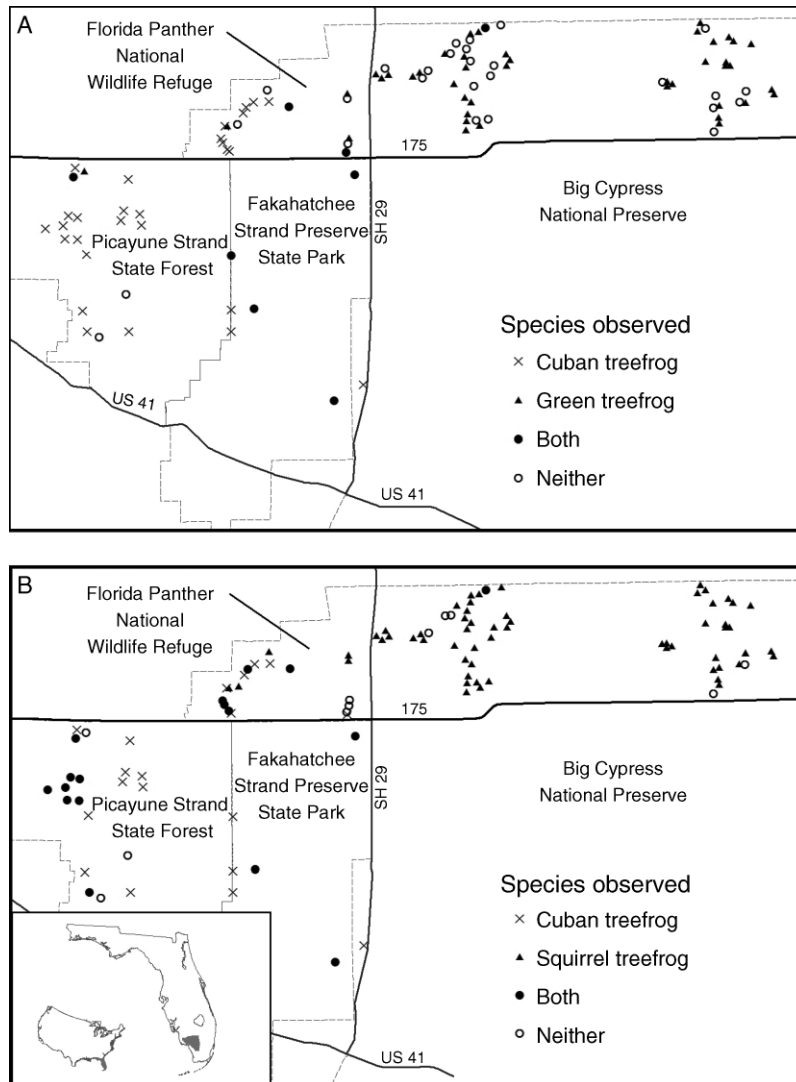


FIG. 2. Maps of sample locations. (A) Locations where green treefrogs and Cuban treefrogs were detected either alone or together. (B) Locations where squirrel treefrogs and Cuban treefrogs were detected either alone or together. The inset map indicates the study area location in southern Florida and the location of Florida in the United States.

County, southwest Florida, USA (Fig. 2). Sites were chosen from randomly selected points within one km of unpaved roads, our primary means of access. Sites were located in the two most abundant habitat types in which treefrogs are normally found in the study area: pine savanna and swamp forest (Madden et al. 1999). Pine savanna sites have a canopy of south Florida slash pine (*Pinus elliottii* var. *densa*) with a shrubby understory dominated by wax myrtle (*Myrica cerifera*) and saw palmetto (*Serenoa repens*). Pine savannas are rarely inundated during the wetter summer months and are maintained with frequent fire (Madden et al. 1999). Swamp forest sites are typically cypress strands (*Taxodium ascendens* and *T. distichum*) but also include mixed hardwood (*Quercus* spp., *Acer rubrum*, and *Fraxinus caroliniana*) and sites dominated by willow

(*Salix caroliniana*) or cabbage palm (*Sabal palmetto*). Swamp forests are usually inundated for three or more months during the wet summer (Duever 2005). Because the original source of Cuban treefrogs in this area is believed to be to the west near Naples, Florida, we used longitude of each sample location, along with habitat type, as a covariate of occurrence of Cuban treefrogs.

Each of the 107 sites was sampled for treefrogs three times during a period of 3–105 days (mean = 24 days) between February 2007 and June 2008. We believe the period between the first and last sample of each site was sufficiently short that demographic closure to changes in occupancy state may safely be assumed for each site. During each visit treefrogs were sampled by at least two experienced observers for one person-hour using a standard, nighttime, visual encounter survey technique

TABLE 1. Comparison of models fit to the detections of Cuban treefrogs, green treefrogs, and squirrel treefrogs.

Model	Cuban treefrog		Native treefrogs		Number of parameters	-log <i>L</i>	BIC	Probability of model
	$\psi$	<i>p</i>	$\psi$	<i>p</i>				
1	hab + lon	temp	pred	temp	13	432.1300	925.01	0.024
2	hab + lon	constant	pred	temp	12	432.1599	920.39	0.246
3	lon	temp	pred	temp	12	433.4797	923.03	0.066
4	lon	constant	pred × hab	temp	15	431.7846	933.66	<0.001
5	lon	constant	pred	pred + temp	13	432.5218	925.79	0.017
6	lon	constant	pred	temp	11	433.5421	918.49	0.638
7	lon	constant	hab	temp	11	450.4996	952.40	<0.001
8	lon	constant	pred	constant	9	442.4524	926.96	0.009

Notes: Models differ by their effects of covariates (hab = habitat, lon = longitude, temp = air temperature, pred = occurrence of Cuban treefrogs) on probabilities of species occurrence  $\psi$  and species detection *p*. The maximized log-likelihood function is denoted by log *L*. Model comparisons are based on the Bayesian information criterion (BIC), which is used to approximate the posterior model probability.

(Crump and Scott 1994). Observers used high-power headlamps to search the area within a circular plot of 20 m radius (1257 m<sup>2</sup>). To be certain of species identifications, attempts were made to capture every individual treefrog that was detected. Body length (snout-to-urostyle distance measured to the nearest mm) was recorded for each captured frog, and basic weather data (air temperature and relative humidity) were recorded during each visit to every sample site.

### Results

To illustrate our model of co-occurrence, we analyzed the detections of Cuban treefrogs and two native species, the green treefrog (*Hyla cinerea*) and the squirrel treefrog (*H. squirella*). During the surveys we observed 193 Cuban treefrogs at 36 of the 107 sample locations (Fig. 2). We observed 171 and 568 green and squirrel treefrogs at 52 and 77 of the sample locations, respectively. The mean length of Cuban treefrogs captured during our sampling was 38.6 mm (SD = 10.9). The mean length of green treefrogs was 30.3 mm (SD = 10.6), and the mean length of squirrel treefrogs was 22.6 mm (SD = 4.9).

We compared eight plausible models of the data using the Bayesian information criterion (BIC) to approximate posterior model probabilities (Table 1). The model with the lowest BIC value (918.49) and the highest posterior probability (0.638) assumes that probabilities of occurrence of the native treefrog species depend only on the occupancy status of Cuban treefrogs and that the probability of occurrence of Cuban treefrogs is a function of longitude. In addition, this model assumes that detection probabilities of native treefrog species increase with air temperature and that the probability of detection of Cuban treefrogs is constant. The model with the second highest posterior probability (0.246) contains an identical set of covariates except that habitat is also included as a covariate of occurrence of Cuban treefrogs.

Estimates of the parameters of the model with highest posterior probability (model 6) imply that occurrence probabilities of green treefrogs and squirrel treefrogs were significantly lower at sites where Cuban treefrogs were present than at sites where this species was absent

(Fig. 3). Using the odds ratio to express these differences, sites without Cuban treefrogs were 9.0 times more likely to be occupied by green treefrogs than sites with Cuban treefrogs (95% CI = 2.3–35.9), and squirrel treefrogs were 15.7 times more likely to occur at sites without Cuban treefrogs (95% CI = 3.8–64.3). Estimates of  $\psi$  for the Cuban treefrog suggest that there was a strong spatial gradient in the occurrence of this species ( $\hat{\beta}_1^B = -4.99$ , SE = 1.20). Sites on the western edge of the study area are estimated to have been nearly 100% occupied, whereas sites on the eastern edge are estimated to have had virtually no occurrences of Cuban treefrogs.

Estimates of the parameters of model 2, which had the second highest posterior probability, are very similar to those of model 6, with one exception. In model 2 estimates of  $\psi$  for Cuban treefrogs were higher in pine savanna habitats than in swamp forest habitats ( $\hat{\beta}_{01}^B = -1.22$ , SE = 0.71 vs.  $\hat{\beta}_{02}^B = -2.91$ , SE = 1.00). However, the spatial

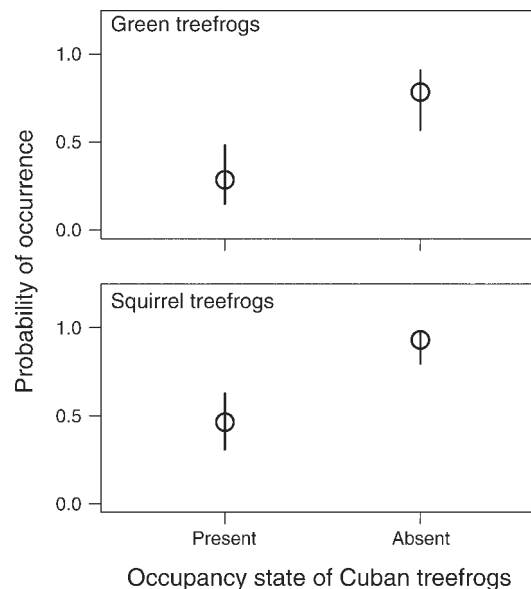


FIG. 3. Estimates of probabilities of occurrence of native treefrog species in the presence and absence of Cuban treefrogs. Error bars indicate 95% confidence intervals.

gradient in Cuban treefrog occurrence estimated using model 2 ( $\beta_1^B = -5.65$ ,  $SE = 1.75$ ) is similar to the estimate obtained using model 6. Qualitatively, the inferences obtained using models 2 and 6 are nearly identical.

#### DISCUSSION

The effect of interspecific interactions on species co-occurrence patterns has long been a topic of interest in ecology. More recently, the potential for negative interactions following the spread of non-native species or following the range expansion of native species (e.g., due to climate change) has fueled additional interest in estimating patterns of species co-occurrence, particularly among applied ecologists. Our approach for estimating co-occurrence of interacting species provides a useful framework wherein the effects of covariates thought to be informative of species occurrence or detectability can be modeled quite naturally while honoring the nature of the interactions among species. We have described this model in terms of predator-prey interactions; however, the model also applies to other types of asymmetric interspecific interactions, such as interference competition (Connell 1983). Our model of co-occurrence can therefore be used in a variety of ecological inference problems.

Our model of co-occurrence can be viewed as a restricted version of the model developed by MacKenzie et al. (2004). In their model the probability that two species ( $A$  and  $B$ ) are both present (i.e., co-occur) is denoted by the parameter  $\psi^{AB}$ . Our approach reparameterizes this model to specify an ecological interaction between the two species wherein one species is assumed to be dominant and the other subordinate. Specifically, we assume that the probability of co-occurrence between species  $A$  and  $B$  is factored into a product of conditional and marginal probabilities of species occurrence:  $\psi^{AB} = \psi^{A|B}\psi^B$ . In addition to formalizing the interaction between species (i.e., the dependence of species  $A$  on occurrence of species  $B$ ), this factorization yields other important benefits. In particular, it allows the effects of spatially varying covariates (e.g., measures of habitat) on probabilities of species occurrence and co-occurrence to be modeled quite naturally. In our factorization, the occurrence probabilities of each species can be modeled independently using different sets of covariates while honoring the nature of the dependence between species induced by their interaction. This benefit also applies to our covariate-based models of detection probabilities.

In contrast to our approach, MacKenzie et al. (2004) proposed the reparameterization,  $\psi^{AB} = \psi^A\psi^B\gamma$  and  $r^{AB} = r^Ar^B\delta$ , wherein the parameters  $\gamma$  and  $\delta$  denote “species interaction factors” for the probabilities of species occurrence and detection, respectively. In this reparameterization the effects of covariates on  $\psi^{AB}$  are specified by assuming separate, logit-scale models for each component of  $\psi^{AB}$  (that is, for  $\psi^A$ ,  $\psi^B$ , and  $\gamma$ ). While the construction of covariate-based models of  $\psi^A$  and  $\psi^B$  seems simple enough, choosing a set of

covariates for modeling  $\gamma$  does not seem straightforward if different covariates are selected for  $\psi^A$  and  $\psi^B$ . For example, should the model of  $\gamma$  include the union or the intersection of the two sets of covariates selected for  $\psi^A$  and  $\psi^B$ ? In addition to these conceptual difficulties, MacKenzie et al. (2004) noted that, in formulating models of  $\gamma$ , the restrictions imposed by interrelationships among the parameters  $\psi^{AB}$ ,  $\psi^A$ , and  $\psi^B$  must be honored. Unfortunately, doing so can cause numerical problems during parameter estimation, as evidenced in their analysis of terrestrial salamander data (MacKenzie et al. 2004).

Our model of species co-occurrence does not suffer from these difficulties. Different covariates can be used to model the occurrence or detection probabilities of each species without fear of estimation problems caused by range restrictions of parameters. Thus, by providing estimates of the effects of covariates on patterns of species co-occurrence, our parameterization can avoid erroneous conclusions that may be reached if such covariates are excluded from the model (MacKenzie et al. 2004). Furthermore, the results of our simulation study indicate that our model yields unbiased estimates of species occurrence probabilities provided sample sizes are sufficiently large. Our simulations revealed that samples of  $n = 50$ – $100$  locations and  $J = 3$  surveys per location can provide accurate estimates of occurrence probabilities if species are easily detected ( $p \geq 0.5$ ). However, higher sample sizes ( $n > 200$ ) are needed to achieve the same accuracy if species are detected with low probability during each survey.

Our model of co-occurrence was useful in analyzing the data collected in surveys of treefrogs in southern Florida. Our analysis revealed that green treefrogs and squirrel treefrogs are much less likely to occur at sites where Cuban treefrogs are present than at sites occupied by this non-native species. Of course, our analysis cannot reveal the mechanisms by which native treefrog species are excluded from sites with Cuban treefrogs, but various lines of natural history information are available to suggest possible causes of the pattern. Studies on the diet of Cuban treefrogs reveal that they are voracious predators with a broad dietary niche (Meshaka 2001, Owen 2005). Cuban treefrogs in southern Florida primarily prey upon invertebrates, such as beetles (Coleoptera) and roaches (Blattaria), but vertebrate animals also have been found as prey, including at least six species of anurans (Meshaka 2001, Maskell et al. 2003). Adult Cuban treefrogs are larger than the adults of the native treefrogs and rates of batrachophagy in Cuban treefrogs as high as 10% have been recorded from specimens collected in Everglades National Park (Meshaka 2001). Therefore, it is likely that some combination of predation and competition for invertebrate prey allows the larger Cuban treefrog to become dominant at sites where it occurs.

If the native treefrogs were exhibiting a behavioral response to the presence of Cuban treefrogs, we would



expect to see a decrease in detectability of native treefrogs at sites where Cuban treefrogs were present. Our comparison of models indicated that the detectability of native treefrog species was not influenced by the presence of Cuban treefrogs (Table 1). This lack of evidence for a behavioral response of native treefrogs to the predatory Cuban treefrog corroborates findings of other studies. Meshaka (2001) examined squirrel treefrog and green treefrog behavior in the presence of Cuban treefrogs. He observed no avoidance of large Cuban treefrogs by the native species and noted that they did not struggle effectively when attacked by Cuban treefrogs. His conclusion was that the native frogs were naive to the predatory threat of Cuban treefrogs. Hoffman (2007) performed an experiment to determine if Cuban treefrogs exclude green or squirrel treefrogs from PVC pipe refugia, but she found no evidence of exclusion or avoidance by the native species. Collectively, this evidence suggests that the native treefrogs do not perceive the Cuban treefrog as a threat, which may make them even more vulnerable as prey.

The geographic pattern of Cuban treefrog occurrence is striking (Fig. 2). The first report of Cuban treefrogs from the city of Naples, Florida (about 22 km west of our westernmost sample location) was by Duellman and Crombie (1970). This invasive species is now commonly observed on the walls of buildings located within housing developments that are rapidly spreading eastward from the city's center. It is unknown if Cuban treefrogs will continue their eastward expansion into the currently unoccupied sites in Big Cypress National Preserve, but our results show that the major habitats that we surveyed are suitable for Cuban treefrogs. Given our results, it seems likely that if Cuban treefrogs continue to expand their range into natural areas in southern Florida, they will have a negative impact on the occurrence of native treefrog species.

#### ACKNOWLEDGMENTS

I. Bartoszek, S. Funk, C. McKee, and G. Richards provided assistance in the field, and J. Barichivich aided many aspects of this study. B. Glorioso created the map. We thank personnel at Florida Panther National Wildlife Refuge, Big Cypress National Preserve, Picayune Strand State Forest, and Fakahatchee Strand Preserve State Park for permission to work on these lands. All handling of animals was conducted under University of Florida IACUC approval number 001-09FTL. This study was funded by the U.S. Geological Survey Greater Everglades Priority Ecosystems Science program. Manuscript preparation was funded by the USGS Amphibian Research and Monitoring Initiative. The manuscript was greatly improved by the review comments of the two anonymous referees.

#### LITERATURE CITED

- Chase, J. M., P. A. Abrams, J. P. Grover, S. Diehl, P. Chesson, R. D. Hold, S. A. Richards, R. M. Nisbet, and T. J. Case. 2002. The interaction between predation and competition: a review and synthesis. *Ecology Letters* 5:302–315.
- Connell, J. H. 1983. On the prevalence and relative importance of interspecific competition: evidence from field experiments. *American Naturalist* 122:661–696.
- Crump, M. L., and N. J. Scott. 1994. Visual encounter surveys. Pages 84–92 in W. R. Heyer, M. A. Donnelly, R. W. McDiarmid, L. C. Hayek, and M. S. Foster, editors. *Measuring and monitoring co-occurrence of interacting species biological diversity: standard methods for amphibians*. Smithsonian Institution Press, Washington, D.C., USA.
- Duellman, W. E., and R. I. Crombie. 1970. *Hyla septentrionalis*. *Catalogue of American Amphibians and Reptiles* 92.1–92.4.
- Duever, M. J. 2005. Big Cypress regional ecosystem conceptual model. *Wetlands* 25:843–853.
- Hoffman, K. E. 2007. Testing the influence of Cuban treefrogs (*Osteopilus septentrionalis*) on native treefrog detection and abundance. Thesis. University of Florida, Gainesville, Florida, USA.
- MacKenzie, D. I., L. L. Bailey, and J. D. Nichols. 2004. Investigating species co-occurrence patterns when species are detected imperfectly. *Journal of Animal Ecology* 73:546–555.
- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83:2248–2255.
- MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. *Occupancy estimation and modeling*. Elsevier, Amsterdam, The Netherlands.
- Madden, M., D. Jones, and L. Vilchek. 1999. Photointerpretation key for the Everglades vegetation classification system. *Photogrammetric Engineering and Remote Sensing* 65:171–177.
- Maskell, A. J., J. H. Waddle, and K. G. Rice. 2003. *Osteopilus septentrionalis*: diet. *Herpetological Review* 34:137.
- Meshaka, W. E. 1996. Vigility and the Florida distribution of the Cuban treefrog (*Osteopilus septentrionalis*). *Herpetological Review* 27:37–40.
- Meshaka, W. E. 2001. *The Cuban treefrog in Florida*. University of Florida Press, Gainesville, Florida, USA.
- Meshaka, W. E., W. F. Loftus, and T. Steiner. 2000. The herpetofauna of Everglades National Park. *Florida Scientist* 63:84–103.
- Owen, J. L. 2005. *The Cuban treefrog (Osteopilus septentrionalis): distribution, diet, and reproduction of an invasive species in the British Virgin Islands*. Thesis. Texas Tech University, Lubbock, Texas, USA.
- Polis, G. A., C. A. Myers, and R. D. Holt. 1989. The ecology and evolution of intraguild predation: potential competitors that eat each other. *Annual Review of Ecology and Systematics* 20:297–330.
- Royle, J. A., and R. M. Dorazio. 2008. *Hierarchical modeling and inference in ecology*. Academic Press, Amsterdam, The Netherlands.
- Smith, K. G. 2005. Effects of nonindigenous tadpoles on native tadpoles in Florida: evidence of competition. *Biological Conservation* 123:433–441.

#### APPENDIX

Maximum-likelihood estimation of parameters of species co-occurrence models (*Ecological Archives* A020-054-A1).

#### SUPPLEMENT

R and WinBUGS code for fitting species co-occurrence models to treefrog data (*Ecological Archives* A020-054-S1).